DOCUMENT RESUME

ED 164 604 TM 008 168

AUTHOR Dryden, Louis J.
TITLE A Model for Evaluating Compensatory Education Programs at the Local Level.
PUB DATE Mar 78
NOTE 23p.; Paper presented at the Annual Meeting of the American Educational Research Association (62nd, Toronto, Ontario, Canada, March 27-31, 1978)

EDRS PRICE MF-$0.83 HC-$1.67 Plus Postage.
DESCRIPTORS Documentation; *Educational Assessment; Educational Objectives; Educational Programs; *Evaluation Methods; *Guidelines; *Models; *Program Evaluation; Theories

ABSTRACT

A model or conceptual framework of educational program evaluation is proposed based upon three major evaluation constructs: documentation, outcome identification, and judgment of outcomes. A series of specific guidelines are presented for the practicing educator in the form of postulates underlying the constructs. (1) Program evaluation documents the nature of the program, as implemented; identifies outcomes; and judges their worth. (2) Program documentation reports on the program characteristics; reduces the probability of evaluating a non-event; facilitates program replication; and considers the physical features and activities of the program. (3) Program objectives specify intended outcomes, in measurable terms, but may not necessarily be stated in advance of the evaluation. The evaluator assists program personnel in writing objectives and identifying priorities. (4) Unintended as well as intended outcomes should be reported; they may be assessed using informal methods. (5) Evaluation begins when a defined entity exists that can be tested and replicated. The evaluator may intervene at any time after a program has been implemented. (6) Judgments concerning the worth of intended or unintended program outcomes can be made in relation to criterion or experiential standards. Further work is needed in the areas of cost analysis, determining unintended outcomes, and setting standards. (GDC)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED164604

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

A MODEL FOR EVALUATING COMPENSATORY EDUCATION PROGRAMS AT THE LOCAL LEVEL

Louis J. Dryden
Educational Research Institute of British Columbia
Vancouver, B.C.

A paper presented at the annual meeting of the
American Educational Research Association,
Toronto, Ontario, March 27 - 31, 1978

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Louis Dryden

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) AND USERS OF THE ERIC SYSTEM."

The information in this paper was written by Louis J. Dryden in his private capacity. No official support or endorsement by the Educational Research Institute of British Columbia is intended or should be inferred.

TM008 168

INTRODUCTION

In recent years, Canadian school district educators, like their American counterparts, have found themselves on the horns of a dilemma. On the one hand the pressure to provide more and more useful evaluations of their programs has increased their need for evaluation models which delineate the potential steps in the evaluation process. On the other hand the models that have been published in education, most of which are similar in content to those described by Scriven, Stake and Stufflebeam, are general models which do not provide the level of specificity required. The major function of these general models is to provide a rationale from which an evaluation can be developed; it is the overall perspective that is important rather than specific details. The practicing educator, faced with this quantum jump from general to specific, is left to find other guides. For the most part, the search has been complicated by a rapid proliferation, in recent years, of evaluation models. This proliferation of models has made the task of learning about evaluation unnecessarily complex. Some of the models have served to confuse rather than enlighten potential users by using diverse terminologies to describe basically the same process. Even a cursory review of contemporary models reveals that many are almost identical; only the words used to describe them are different.

The author faced this same dilemma when asked to conduct an evaluation of a compensatory education project in a northern Alberta rural school district.* What follows is a description of the construction of a model of the evaluation process which functioned as the conceptual framework for the compensatory evaluation. In constructing the model an attempt was made to avoid the lack of specificity inherent in more general models and to focus on the complementary nature of these models. What is presented in this paper represents an attempt to bridge the gap between conceptualization and

*See Dryden, L.J. and Nygaard, M.H. *A Summative Evaluation Report of the High Prairie School Division No. 48 Compensatory Education Program.* Alberta Evaluation Research Associates Ltd. June 30, 1976.

practice in educational program evaluation. The end product of this endeavor is a set of field-oriented guidelines which can be used by practicing educators who are responsible for evaluating programs, but who have limited theoretical background in the area of evaluation, and little time in which to study the proliferation of ideas and approaches presented in contemporary literature. The guidelines are presented in the form of a number of postulates which underlie the aforementioned model.

The organization of this paper mirrors the process through which the resulting conceptual model of the evaluation process was derived:

- A modelling approach to evaluation theory development
- The purpose of the model
- The construction of the model
- A schematic of the model
- The potential steps in the evaluation process
- Implications of the model for future applications.

A MODELLING APPROACH TO EVALUATION THEORY DEVELOPMENT

Most research endeavors are guided by theory. The theory may be stated explicitly or merely implied by the researcher. His theory may be a vague hunch, an informed guess, a set of inconsistent assumptions or a logically structured explanation. Research is a sequential process. Each successive step is directed largely by the knowledge gained in previous steps and is related to the whole by theory. Like research, evaluation is a sequential process. A lack of adequate theory, however, has been, and still remains, a major limitation in the field of educational program evaluation. Perhaps because evaluation is a practice-oriented process generally conducted in a field setting, evaluators have assumed that theory has little of significance to contribute to their particular endeavors. This assumption is completely false. Theory and conceptualization are not foreign to practice. Administrators and practitioners daily work on the basis of certain underlying

assumptions about the nature and causes of problems and about appropriate intervention methods. An analysis of how decisions are made will reveal that decisions are usually based on a set of theoretical postulates, even though these are likely to be implicit and untested.

Whether explicit or implicit, research and practice are both based on theory. Educational evaluation, which relies on research for design and methodology and practice for its existence, has been found wanting in theory development.

Marx (1963) has identified four modes of theory construction: the use of a model, the deductive procedure, the functional procedure, and the inductive procedure. The four categories are *not* mutually exclusive. He has defined a model as any conceptual analogue, generally of a physical or mathematical nature, which is used to suggest empirical research. According to him, the flow of influence in a model is entirely from the conceptual to the empirical level. The direction of the flow of influence is signficant because once a particular model has been selected, the researcher is not concerned at all about modifying the model itself on the basis of the data obtained by means of it; this insensitivity to data is a major difference between the model and other forms of theory development. A model approach to theory development, therefore, can be chosen for its heuristic or tool value and in this mode it can be used to focus the research effort.

Other individuals writing on the nature and use of models in science (Lachman, 1960: Chapanis, 1961: Kaplan, 1964: Black, 1973) have noted the multiplicity of descriptions of the term. Chapanis' (1961) comments are particularly pertinent:

> Models, we find from reading the literature, come in a bewildering variety. There are mechanical models, true models, adequate models, distorted models, dissimilar models, static models, dynamic models, structural models, iconic models, analogic models, symbolic models, material models, mathematical models, and analytic models - just to name a few. (p. 109)

Given such diverse descriptions of models it is somewhat of an understatement that the term model can be defined in many ways. In an attempt at clarifica-

tion, Chapanis (1961) has adopted the following simple definition of the term model:

> Scientific or engineering models are representations, or likenesses, of certain aspects of complex events, structures, or systems, made by using symbols or objects which in some way resemble the thing being modeled. (p. 109)

His definition serves to identify two basic kinds of models: replica models and symbolic models. According to Chapanis the essential thing about replica models is that they look like the thing being modelled in some respect. On the other hand, symbolic models are intangible in the sense that they make use of ideas, concepts, and abstract symbols to represent the objects being modelled.

Black (1973) has delimited his discussion of models to four categories: (1) scale models, (2) analogue models in which there is an isomorphic relationship between the model and the original entity, (3) exemplar models which outline designs or conceptions of things worthy of imitation and (4) theoretical models in which one uses an area better understood in order to solve a particular problem. It is argued in this paper that the term model, as typically *used* in educational evaluation appears to refer to the third definition above, that is, they are simply exemplary designs or procedures and are probably best viewed as descriptive conceptualizations of evaluation processes. It is further argued that educational evaluation models *are* symbolic in nature, that is, they are intangible in the sense that they make use of ideas, concepts and abstract symbols to represent the evaluation process.

Borich (1974) indicates that evaluators have probably chosen to use exemplar models by design and not by chance:

> Because of the complexity of events the evaluator must describe, he is usually cautious about using a model that is too specific to a set of phenomena. Even though the evaluator attempts to be precise, there are often so many inputs at work in a given program or product that it is difficult to describe all of them with a great deal of precision. (p. 145)

He also maintains that:

> In contrast to the scientist, the evaluator usually spends much less

> time verifying his model than he spends using it. The evaluator sees his models as a means to an end; if it works reasonably well in achieving that end, the evaluator rarely takes the time to document the effectiveness of the tool that got him there. (Borich, 1974, p. 145).

His second comment closely resembles Marx's (1963) description of the use of a model approach to theory development.

Two more important points need to be clarified. The first concerns the distinction between a theory and a model. Considerable confusion exists between the two terms. This situation probably arises, in part, because the words theory and model have so many different connotations. A useful description of the difference between a theory and a model is supplied by Chapanis (1961):

> A model is an analogy and no one expects an analogy to be completely accurate. ... With a model it is not even important that any of these conditions be true, for a model can tolerate a considerable amount of slop. It is only an analogy, a statement that in some ways the thing modelled behaves "like this". ... Theory, on the other hand, is a conceptual system which attempts to describe the real thing. The basic elements, or pieces, of a theory are actually supposed to be there in the thing about which you are theorizing and they are supposed to behave the way the theory says. Whereas a model tolerates some facts which are clearly not in accord with it, facts which do not agree with theory are fatal to the theory. (p. 112)

Models, in a word, are judged by criteria of usefulness; theories, by criteria of truthfulness.

The second point concerns the tendency of model builders to reify their products. The need to be aware of this tendency is reinforced by Chapanis (1961). He considers that the worst error committed in the name of models is to forget that at best a model represents only a *part* - and usually only a *small* part - of the thing being modelled.

Obviously, in theory development one must have an understanding of the situation in general before one can formalize theories. MacIver and Holdaway (1966) have speculated that much of the hit and miss in research today is due to the lack of comprehensive conceptual models. They maintain that theory development starts with a conceptual model of some area of knowledge. The construction of a conceptual model,therefore, precedes the

development of theory.

THE PURPOSE OF THE MODEL

A model is a methodological device which can be used to develop theory incrementally. A model can also be used to guide and focus empirical enquiry. For the evaluator, a model is symbolic and is usually of the exemplar type. It represents a desirable descriptive conceptualization of the evaluation process. The use of an exemplar model also affords the evaluator the flexibility needed when working in the "natural" laboratory. Therefore, until sufficient evidence has accumulated upon which to build a theory of evaluation, it is assumed that evaluations will be facilitated by the development and use of symbolic exemplar models which describe the potential steps in the evaluation process.

To summarize, the purposes of a model of the program evaluation process are:

- To serve as a theoretical base
- To identify patterns of relationships between variables thereby focusing the evaluation
- To identify elements that are important thereby accounting for what otherwise might be overlooked
- To explicate the potential steps in the evaluation process
- To better explain the processes involved in an effective evaluation
- To avoid problems inherent in the use of general evaluation models
- To simplify the evaluation process

THE CONSTRUCTION OF THE MODEL

A conceptual model of the program evaluation process was constructed by abstracting selected concepts (obtained from a review of the works of writers on educational evaluation and other areas within the field of applied social research) and formulating them into postulates. These postulates were, in turn, used as the basis of the model.

A Definition of Program Evaluation

To start with, in the development of a model of the program evaluation process, it was necessary to derive a suitable definition of program evaluation. Most would agree that a major failing of evaluation today stems from the lack of adequate definitions.

Generally, variations in definitions of evaluation reflect fundamental differences in philosophical and methodological positions. Those who equate evaluation with measurement emphasize determining whether program objectives (usually stated in behavioral terms) have been achieved. Tyler's work is perhaps most representative of this approach. Those who equate evaluation with decision making perceive evaluation as a service, an aid to decision makers. Cronbach, Alkin and Stufflebeam are representative of this approach. Those who equate evaluation with judgment stress the worth of the work that has been accomplished. Scriven's (1967) position is that there is *no* evaluation until judgment has been rendered; the *goal* of evaluation is always the same, to establish the merit or worth of something. Although hesitant to stress evaluator judgment, Stake (1967) nonetheless suggests that description and judgment are the two basic acts of evaluation. Hemphill (1973), representing an educational research and development point of view, has also emphasized the judgmental aspect of evaluation. According to him, educational evaluation produces evidence that can sharpen judgments about given states of affairs in education; evaluation always implies judgment.

The trend in educational evaluation is towards judging the worth of something. The purpose of evaluation may, therefore, be expressed in terms of ascertaining the effectiveness, the merit, the success of the phenomenon under study. It is argued therefore that a definition of educational program evaluation should include the judgment of worth. Before one can evaluate something, however, the nature of that something must be known. Therefore it is argued that describing the nature of the program to be evaluated is also essential to a definition of program evaluation. Given that these two assumptions are valid, the question arises: "Upon what does one pass judgment?" A general answer is supplied by Stake (1967). He has stated that judgmental data are collected about antecedents, transactions

and outcomes. A more specific answer is provided by Wholey (1972). Representing a policy science perspective, Wholey has suggested that the essence of program evaluation, is the assessment of program outcome - what happened that would not have happened in the absence of the program. Thus, it is further argued that a definition of evaluation should emphasize judging the worth of program outcomes.

Together, these three assumptions form the basis of the definition of evaluation used in this paper. The definition is stated as the first postulate and is elemental to the construction of the model of the program evaluation process which follows:

Postulate 1:

Educational program evaluation is the process of documenting the nature of a given program as it has been actually implemented; identifying intended and/or unintended program outcomes; and, judging the worth of these outcomes at a given point in time.

Constructs Proposed

The above definition contains three major evaluation constructs: documentation, identification, and judgment. A description of and the rationale underlying each construct is presented below.

Documentation

Central to the above definition is the notion that the purpose of program evaluation is to ascertain the worth of a program as *it has been actually implemented.* To make this determination, the program to be evaluated must be extensively and thoroughly described. This activity is referred to as program documentation and is defined as a report of a program's characteristics: its participants, materials, and activities.

Several factors indicate that program documentation is one of the most important activities in program evaluation. Unless the evaluator carefully documents the activities that comprise a project, there is no way to tell whether, in fact, a project was actually implemented. This question is raised by Wrightstone (1969). He asserts that it may well be that the principal, or at least a major explanatory variable in the effectiveness or

ineffectiveness of a given program is the loss of input between its dissemination by authorities and its reception by the child it is intended to effect. Similarly, many evaluations report findings of no significant difference between the results obtained from the treatment and non-treatment groups. Perhaps, there was no difference between the treatments being compared. A consequence, in this instance, of failing to document thoroughly a project is described by Charters and Jones (1973):

> What is *not* standard practice in evaluation studies is to describe, let alone measure, how the programs in "experimental" and "control" situations actually differ from one another - or even to certify that they do. A serious consequence of slighting the alleged "independent variable" is that elaborately designed evaluation studies may sometimes end up appraising non-events, with no one the wiser. (p. 5)

The two summative components of the model developed at the Centre for the Study of Evaluation (Alkin and Fitz-Gibbon, 1975), are documentation and outcomes evaluation. The model provides a useful interpretation of the documentation process. According to the CSE model, documentation, whether brief or detailed, should mention some facts about the program's *physical features* and program *activities*. In documenting the physical features reference should be made to program staff (including administrators, teachers, aides, parent volunteers, etc.), resources (primarily equipment and materials), school and classroom organizational structure and students. Activities are, by definition, time-and-people-related events which are intended to occur as necessary parts of a program. In documenting the activities of a program, reference should be made to materials used and to descriptions of activities which take place in the program.

In summary, then, it may be stated that:

Postulate 2:

Documentation of a program is a report of that program's characteristics: its participants, materials, and activities.

Postulate 3:

Program documentation reduces the probability of evaluating a non-event.

Postulate 4:

Documentation facilities program replication and/or expansion.

Postulate 5:

Documentation includes the consideration of a program's physical features and its activities.

Identification

The second construct embodies the identification of program outcomes. Program outcomes may be intended and unintended, short-term or long-term. Intended program outcomes can be identified by examining the objectives of a program. Program objectives which are written in measurable terms are referred to as performance objectives. Such objectives are characterized by a precise statement of a goal (the intent of the activity) and, in addition, specifies some type of criterion standard that can be used to determine the level of success obtained in achieving that stated goal. Thus, a well-written performance objective will specify (1) an intended program outcome, and (2) a criterion standard for ascertaining the degree to which an intended outcome has been achieved. In addition, program objectives which are stated in performance terms will effectively prescribe evaluation procedures. By stating the objectives in performance terms, the criterion (outcome) variables to be measured are identified. The evaluator must then locate information sources, develop an appropriate investigatory framework for collecting the information, and select and/or develop instrumentation for gathering the needed information.

In summary, then, it may be stated that:

Postulate 6:

Program objectives specify intended program outcomes.

Postulate 7:

Program objectives which are stated in measurable (performance) terms are to be preferred to those which are non-measurable.

Program objectives, as noted above, supply the evaluator with important information and, excluding Scriven's goal-free approach, are an integral part of most comprehensive evaluations. It cannot be assumed, however, that

the objectives of a program which one wishes to evaluate are known or stated in advance of the evaluation. If very little program planning preceded implementation, then an evaluator may very well be asked to evaluate a program for which no clear statement of objectives has been made. In addition, programs change over time and objectives likewise, change. Accurate and fair evaluations must take these two circumstances into account. Therefore, the first major task, following program documentation, is to make the program's objectives explicit, that is, the evaluator constructs a list of program objectives that explain without ambiguity what the planners of the program agree it should accomplish. Support for this notion is provided by Worthen (1974) and Freeman and Sherwood (1965). Worthen maintains that an important activity in getting many evaluations under way is eliciting from the responsible persons the goals and objectives of the program. Similarly, Freeman and Sherwood believe that if the practitioner or the administrator does not or cannot specify both the objectives and lying assumptions, it will remain for the researcher to infer these from the components of the program or insistently to help his program colleagues state the goals and their theoretical links.

Worthen and Sanders (1973) argue that not only must the evaluator help to identify program goals and objectives but (s)he must assist program personnel to rank order program goals and objectives. In order to do this it becomes necessary to assign various weights to the different objectives of a program. Few writers in the field of educational evaluation have considered this aspect and strategies for accomplishing this end are almost nonexistent. Another related question concerns the task of stating program objectives in performance terms. Stake (1967) and Worthen (1974) indicate that it is up to the evaluator to assist in writing performance objectives. Worthen, however, provides a caveat:

> The onus on making this translation (translating broad objectives into specific measurable objectives) clearly lies with the evaluator who possesses the technical skill for doing so and not on program personnel to whom the language of operationalism and behaviorism is foreign and unfamiliar. Of course, the translated objectives must be reviewed by program personnel to prevent unconscious biases of the evaluator from producing operational objectives different in intent from the broad objectives with which the program began. (p. 15)

In summary, then, it may be stated that:

Postulate 8:

It should not be assumed that the objectives of a program which one wishes to evaluate are stated in advance of the evaluation.

Postulate 9:

The evaluator must assist program personnel to identify and rank order program goals and objectives.

Postulate 10:

The evaluator assists program personnel in writing measurable objectives.

A major limitation in focusing only upon program objectives (intended outcomes), is that the evaluator may miss unanticipated side-effects (unintended outcomes). Such side-effects may be more significant than the planned-for main-effects. The evaluator who makes a conscious attempt to detect project side-effects will be attentive to a wider range of outcomes than might be the case of the evaluator who looks only for results consonant with project objectives. At present, according to Deniston et al.(1968), there is no systematic procedure for assessing any unplanned impact of activities although activities performed for a specific purpose may indeed have side-effects. They recommend that program personnel attempt to assess side-effects on a subjective, impressionistic basis until a more systematic way of measuring them is developed. Bloom (1969) also recommends a relatively informal evaluation procedure for assessing program side-effects. According to him it is quite likely that many of the unintended positive as well as negative outcomes of instruction can be detected by competent observers who approach the problem in a clinical way.

In summary, then, it may be stated that:

Postulate 11:

The effectiveness of programs should be reported in terms of the worth of unintended as well as intended program outcomes.

Postulate 12:

In assessing any unplanned impact of program activities the evaluator must rely on relatively informal methods including the subjective,

impressionistic observations of persons involved in or affected by the program to be evaluated.

A final question that must be addressed concerns the point of intervention of the evaluator. Alkin (1969) and Stufflebeam et al. (1971) maintain that the evaluator intervenes early in the program development process, that is, (s)he assists with both needs assessment and program planning. Several writers in the field of educational evaluation (Worthen and Sanders, 1973; Lange, 1974) have argued that activities such as needs assessment and program planning are non-evaluative. Additional evidence in support of these claims is provided by Weiss (1972). She maintains that a analysis of the system can help program developers as they seek direction but it is not evaluation. Evaluation comes later, when there is an entity, however complex and interrelated, that can be defined, tested and replicated. Similarly, Guttentag (1973) maintains that the outcomes of planning is the first stage in the evaluation research process. The point made by Weiss is crucial to understanding the definition of program evaluation used here. According to the definition, the evaluator focuses on outcomes. Outcomes are the consequences of implementing a defined program. The evaluator may, therefore, intervene at any time after a program has been implemented. If the evaluator is consulted shortly after the implementation of a defined program has commenced he would focus on immediate or short-range outcomes and function in a formative role. Again, the outcomes would be intended and unintended.

In summary, then, it may be stated that:

Postulate 13:

The evaluation process commences when a defined entity exists.

Postulate 14:

The evaluator may intervene at any time after a program has been implemented.

Judgment

The third construct contained in the definition used here concerns the judgment of program outcomes. The evaluator's final task, according to

Scriven (1967), is to formulate an overall assessment of the worth of the program. Stake (1970) maintains that judgment data help to show that the design of a program makes sense, or that it does not. Without judgment data an evaluator cannot show that a program is succeeding. The major distinguishing characteristic of evaluation is that it requires judgment.

Judgments are made about standards of performance. Whenever we evaluate something we are involved in making comparisons. We compare the object that we are evaluating against something else which we call a standard. A standard comprises qualities, characteristics, behaviors, conditions, etc., which are considered desirable levels or qualities of performance. Judgments about the worth of an object will be based on the extent to which one is able to determine that the object meets the characteristics of the standards. Stake (1967) maintains that before an evaluation can be completed, there must be judgment in relation to two types of standards. The first, absolute standards, can be in the form of arbitrary levels on achievement tests and the second, relative standards, exist in the form of descriptions of other programs which attempted to obtain similar objectives. He carefully distinguishes between the act of setting standards and the act of judging. He admits that the latter is a relatively subjective commitment in evaluation. In a general discussion of criteria of judgment concerning innovations, Collier (1972) provides support for Stake's observation:

> We cannot divest ourselves of our subjective presuppositions; we cannot detach ourselves from subjective factors. *The search for purely objective criteria in education is an illusion.* (p.181)

Two types of standards, not unlike Stake's (1967) are contained in the model. The first, *criterion* (absolute) *standards* are those standards inherent in performance objectives. Criterion standards can be utilized to determine the degree to which an intended outcome has been achieved. This means that one indice of the worth of a program is a function of the extent to which the intended outcomes are reached and, therefore, represents the standard against which the program will be judged.

While the objectives per se may not be evaluated under this approach (that is, the objectives may not be examined for their intrinsic worth), the standards employed are at least clear and certainly related to the inten-

tations of the program. However, it is also true that building the standard against which the program will be compared around the intended outcomes of the program fails to incorporate into the standard the full range of meaning that makes the program what it is in the experience of its members. This can lead to the feeling, at least on the part of program staff, that the standard against which the program is judged does not take into account sufficient detail about the program and thus judges it too superficially. There may be, in addition, the feeling that the operational definitions agreed to do not convey enough of the context of the program and are thus too simplistic.

In an attempt to overcome the problems inherent in using only criterion standards as the single indice of program worth, a second set of standards was considered. The second set, *experiential standards* refer to those standards of worth inherent in the value structure of those persons involved in, or affected by the program to be evaluated. Experiential standards represent the summative judgments that people make about the overall program or some component of it. Such standards are comparative in nature and are based on experience. Judgments, based on experiential standards, can be made about the worth of intended and unintended program outcomes. Stake (1970) considers that the summative judgments that people make about a program are perhaps the most critical to good evaluation.

In summary, then, it may be stated that:

Postulate 15:

The purpose of evaluation may be expressed in terms of ascertaining the effectiveness, the merit, the success of the phenomenon under study.

Postulate 16:

Judgments about the worth of intended and unintended program outcomes can be made in relation to criterion and/or experiential standards.

A SCHEMATIC OF THE MODEL

The above three categories of postulates can be logically arranged to form a schematic diagram representative of a conceptual model of the

program evaluation process. This diagram, represented by Figure 1, depicts

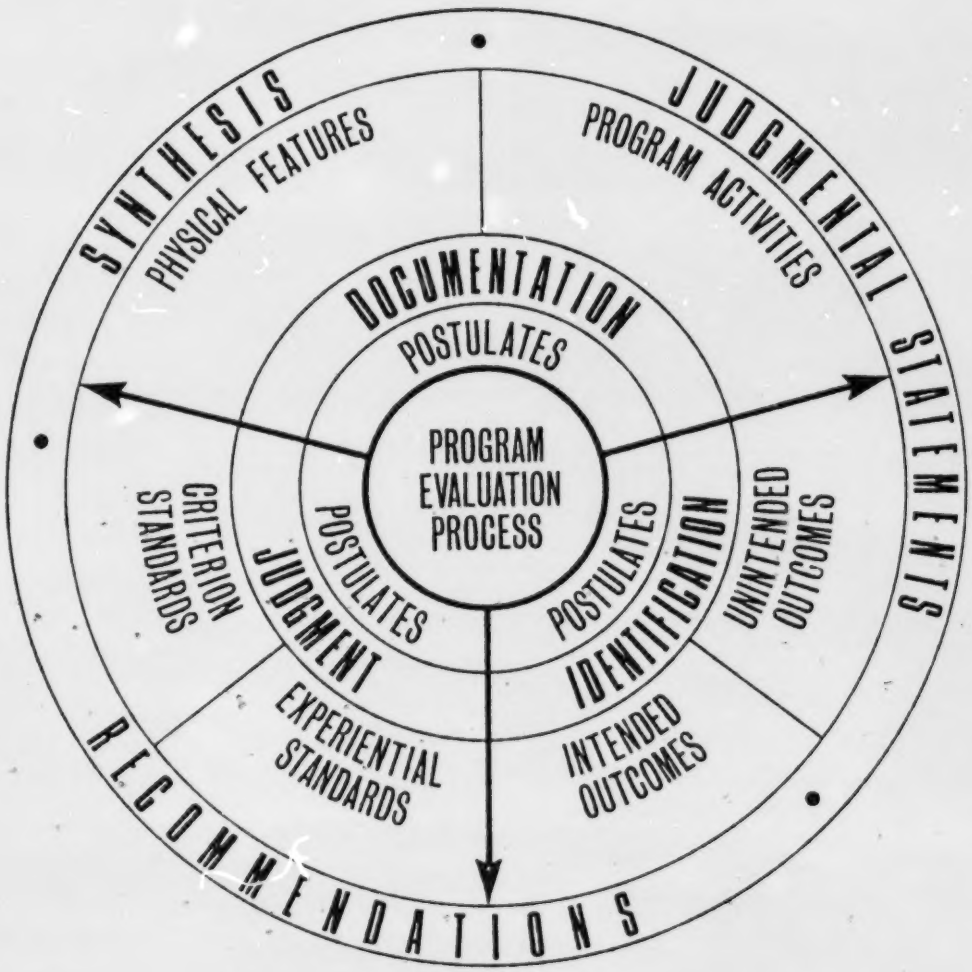

Figure 1. A conceptual model of the program evaluation process.

in schematic form the interrelationships between the three major constructs: documentation, identification, and judgment.

POTENTIAL STEPS IN THE EVALUATION PROCESS

As evidenced in Figure 1, a different set of evaluative activities is associated with each construct. In all, nine evaluative activities are subsumed within the three basic constructs:

Document the program

- Describe the program's physical features. Reference should be made to program staff (including administrators, teachers. aides, parent volunteers, etc.), resources (primarily equipment, materials and facilities), school and classroom organizational structures and students.
- Describe the activities of the program as it is actually implemented. Reference should be made to time and people related events including materials used.

Identify program outcomes

- Identify and rank order the goals and objectives (intended outcomes) of the program. Do not assume that the goals and objectives will be stated in advance of the evaluation.
- Formulate the program's objectives into measurable (performance terms). The evaluator assists program personnel in this activity.
- Identify, in part, the unintended program outcomes by informal methods including the subjective, impressionistic observations of those persons involved in or affected by the program.

Judge program outcomes

- Compare the intended outcomes with the criterion standards contained in the program's performance objectives
- Elicit the judgments of those persons involved in or affected the program.

Synthesize data

- Synthesize all available data into judgmental statements

Make recommendations

- Formulate and present recommendations to client group(s)

IMPLICATIONS OF THE MODEL FOR FUTURE APPLICATIONS

On the basis of experience gained in the use of the model in evaluating the aforementioned compensatory education project and its subsequent use in

two additional evaluation studies,* it appears that the model is, indeed a useful heuristic device for guiding evaluations of educational programs. The author has used the nine evaluation activities inherent in the model and the guidelines as described by the postulates underlying the model for conducting evaluations in three rather diverse contexts.

The model provides a systematic way of arranging descriptive and judgmental data and emphasizes the interrelationships between them. It requires specification of performance objectives thereby focusing both evaluative and program development efforts and considers both criterion and experiential standards in formulating judgments. Limitations include inadequate methods for determining unintended program outcomes or side-effects and for establishing standards. The model also neglects cost considerations.

The ultimate test of the model in terms of its value to the practicing school district educator is the usefulness of the data it provides. The model is, therefore, in need of further field testing in other school systems and in other contexts in order to determine the validity and usefulness of the results produced.

*See Dryden, L.J. *A Report of An Evaluation of An Adult Education French Language Program.* Educational Research Institute of British Columbia. January 30, 1978 and Dryden, L.J. *A Report of An Evaluation of An ESL-TV Project.* Educational Research Institute of British Columbia, (in progress).

BIBLIOGRAPHY

Alkin, M.C. Evaluation theory development. *Evaluation Comment*, 1969, *2* (1), 2-7.

Alkin, M.C. and Fitz-Gibbon, A. Methods and theories of evaluating programs. *Journal of Research and Development in Education*, 1975, *9* (3), 2-15.

Black, M. Models and archetypes. In H.S. Broudy, R.H. Ennis, and L.I. Krimerman (Eds.), *Philosophy of Educational Research*. New York: John Wiley and Sons, 1973.

Bloom, B.S. Some theoretical issues relating to educational evaluation. In R.W. Tyler (Ed.), *Educational Evaluation: New Roles New Means*. Sixty-eighth Yearbook of the National Society for the Study of Education, Part II. Chicago: University of Chicago Press, 1969.

Borich, G.D. (Ed.). *Evaluating Educational Programs and Products*. Englewood Cliffs, N.J.: Educational Technology, 1974.

Cain, C.G. and Hollister, R.G. The methodology of evaluating social action programs. In P.H. Rossi and W. Williams (Eds.), *Evaluating Social Action Programs: Theory, Practice and Politics*. New York: Seminar Press, 1972.

Chapanis, A. Man, machines and models. In M.H. Marx (Ed.), *Theories in Contemproary Psychology*. New York: Macmillan, 1963.

Charters, W.W., Jr., and Jones, J.E. On the risk of appraising non-events in program evaluation. *Educational Researcher*, 1973, 2 (11), 5-7.

Collier, K.G. Innovation: criteria of judgment. *British Journal of Educational Technology*, 1972, *3* (3), 175-185.

Cronbach, L.J. Course improvement through evaluation. *Teachers College Record*, 1963, *64* (8), 672-683.

Deniston, O.L., Rosenstock, I.M., and Getting, V.A. Evaluation of program effectiveness. *Public Health Reports*, 1968, *33* (4), 323-335.

Freeman, H.E., and Sherwood, C.C. Research in large-scale intervention programs. *Journal of Social Issues*, 1965, *21* (1), 11-27.

Guttentag, M. Subjectivitiy and its use in evaluation research. *Evaluation*, 1973, *1* (2), 60-65.

Hemphill, J.K., and Rosenau, F.S. (Eds.) *Educational Development: a new discipline for self renewal*. Eugene, Oregon: Centre for the Advanced Study of Educational Administration, 1973.

Kaplan, A. *The conduct of inquiry; methodology for behavioral science.* San Francisco, Ca.: Chandler, 1964.

Lachman, R. The model in theory construction. In M.H. Marx (Ed.). *Theories in Contemporary Psychology.* New York: Macmillan, 1963.

Lange, R.R. A search for utility in new evaluation thought. *Theory Into Practice,* 1974, *13* (1), 22-30.

MacIvor, D.A., and Holdaway, E.A. An analysis of the use of models in education. *Alberta Journal of Educational Research,* 1966, *12* (3), 163-187.

Marx, M.H. (Ed.). *Theories in Contemporary Psychology.* New York: Macmillan, 1963.

Popham, W.J. *Educational Evaluation.* Englewood Cliffs, N.J.: Prentice-Hall, 1975.

Scriven M. The methodology of evaluation. *Perspectives of Curriculum Evaluation.* AERA Monograph Series on Curriculum Evaluation, No. 1. Chicago, Ill.: Rank McNally, 1967

Scriven M. Prose and cons about goal-free evaluation. *Evaluation Comments,* 1972, *3* (4), 36-38.

Smith, E.R. and Tyler, R.W. *Appraising and Recording Student Progress.* New York: Harper and Row, 1942.

Stake, R.E. The coutenance of educational evaluation. *Teachers College Record,* 1967, *68* (7), 523-540.

Stake, R.E. Objectives, priorities, and other judgment data. *Review of Educational Research,* 1970, *40* (2), 181-213.

Stufflebeam, D.L., Foley, WlJ., Gephart, W.J., Guba, E.G., Hammond, R.L., Merriam, H.O., and Provus, M.M. *Educational Evaluation and Decision Making.* Itasca, Ill.: F.E. Peacock, 1971.

Tyler, R.W. General statement on evaluation. *Journal of Educational Research,* 1942, *35* (7), 492-501.

Weiss, C.H. *Evaluation Research: Methods of Assessing Program Effectiveness.* Englewood Cliffs, N.J.: Prentice-Hall, 1972.

Wholey, J.S. What can we actually get from program effectiveness? *Policy Sciences,* 1972, *3* (3), 361-369.

Worthen, B.L., and Sanders, J.R. *Educational Evaluation: Theory and Practice.* Worthington, Ohio: Charles A. Jones, 1973.

Worthen, B.R. Some tasks and competencies frequently required in educational research and evaluation. Occasional Paper #5. Portland, Oregon: Office of Marketing and Dissemination, Northwest Regional Educational Laboratory, 1976.

Wrightstone, J.W. Evaluating educational programs: a symposium. *The Urban Review,* 1969, *3* (4), 15-17.